

(19)



Europäisches Patentamt

European Patent Office

Office européen des brevets



(11)

EP 0 895 192 A2

(12)

EUROPEAN PATENT APPLICATION

(43) Date of publication:

03.02.1999 Bulletin 1999/05

(51) Int. Cl.⁶: G06T 11/00

(21) Application number: 98114093.2

(22) Date of filing: 28.07.1998

(84) Designated Contracting States:

AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE

Designated Extension States:

AL LT LV MK RO SI

(30) Priority: 28.07.1997 US 901391

(71) Applicant:

DIGITAL EQUIPMENT CORPORATION
Maynard, Massachusetts 01754 (US)

(72) Inventor: Kang, Sing Bing

Cambridge, MA 02138 (US)

(74) Representative: Betten & Resch

Reichenbachstrasse 19
80469 München (DE)(54) **Method for reconstructing a three-dimensional object from a closed-loop sequence of images**

(57) In a computerized method, the three-dimensional structure of an object is recovered from a closed-loop sequence of two-dimensional images taken by a camera undergoing some arbitrary motion. In one type of motion, the camera is held fixed, while the object completes a full 360° rotation about an arbitrary axis. Alternatively, the camera can make a complete rotation about the object. In the sequence of images, feature tracking points are selected using pair-wise image registration. Ellipses are fitted to the feature tracking points

to estimate the tilt of the axis of rotation. A set of variables are set to fixed values while minimizing an image-based objective function to extract a set of first structure and motion parameters. Then the set of variables freed while minimizing of the objective function continues to extract a second set of structure and motion parameters that are substantially the same as the first set of structure and motion parameters.

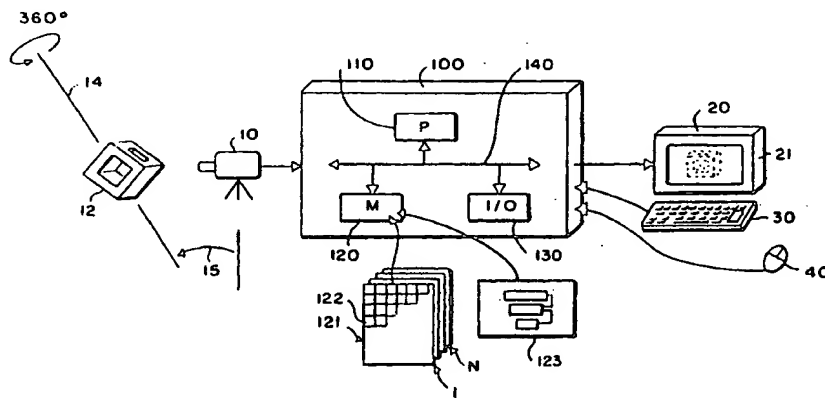


FIG. 1

EP 0 895 192 A2

BEST AVAILABLE COPY

Description

FIELD OF THE INVENTION

[0001] The present invention relates generally to the field of image processing, and more particularly to reconstructing three-dimensional image structure of an object from camera images taken of the object.

BACKGROUND OF THE INVENTION

[0002] In one computerized image processing application, the three-dimensional (3-D) shape of a structure can be recovered from a sequence of digitized images taken by a camera. Typically, one of the first steps taken prior to the actual 3-D reconstruction of the object is usually a process of camera calibration.

[0003] Therefore, particularly germane to 3-D reconstruction using an uncalibrated camera is self-calibration. Self-calibration refers to a recovery of camera parameters based only on correspondences of images taken at different object (or camera) poses. Most work done on self-calibration relies on knowing the motion of the camera (or object), such as pure translational motion, known camera rotation angles, and pure rotation about the center of the camera.

[0004] The traditional approach to 3-D reconstruction with multiple images using an uncalibrated camera apply to affine and projective reconstruction techniques. The traditional approach to reconstruct scaled Euclidean structure is usually from known camera parameters.

[0005] For example, some techniques recover the Euclidean structure from structure rotation or equivalently, camera orbital motion under the assumption that the camera parameters and structure motion are known. Other techniques use a grided annulus pattern inside which the structure is placed. Camera parameters are extracted from the detected pattern, and the 3-D shape of the structure is recovered from silhouettes. Other known techniques use a closed-form solution for scaled Euclidean reconstruction with known intrinsic camera parameters but unknown extrinsic camera parameters. However, these techniques assume the existence of four coplanar correspondences that are not necessarily collinear.

[0006] Recently, 3-D scaled Euclidean structures have been reconstructed from images using an uncalibrated camera. For example, one method reconstructs a scaled Euclidean object under constant but unknown intrinsic camera parameters. There, a minimum of three images are required to recover a unique solution to the intrinsic camera parameters and the Euclidean structure.

[0007] The method is based on recovering an intermediate projective structure. Then, an optimization formulation that is based on the Frobenius norm of a matrix is applied. However, this is not equivalent to the more opti-

mal metric of minimizing feature location errors in a 2-D image space. This method has been extended to show that scaled Euclidean reconstruction under a known image aspect ratio and skew, but varying and unknown focal length and principal point is also possible.

[0008] The assumption there is that the camera is undergoing general motion. It has been shown that it is not possible to reconstruct a scaled Euclidean structure under constrained motion such as pure translational or orbital motion.

[0009] In a two-step approach, a scaled Euclidean structure is recovered from a multiple image sequence with unknown but constant intrinsic parameters. The first stage involves affine camera parameter recovery using the so-called modulus constraint. This is followed in a second stage by conversion to scaled Euclidean structure. This approach can be extended to remove the assumption of a fixed camera focal length.

[0010] One known technique provides a method for camera self-calibration from several views. There a two-step approach is also used to recover a scaled Euclidean structure. The method first recovers projective structure before applying a heuristic search to extract the five intrinsic camera parameters. The heuristic search involves iterating over several sets of initialization values and periodically checking for convergence.

[0011] A detailed characterization of critical motion sequences (CMS) has been given to account for ambiguities in reconstructing a 3-D structure. A critical motion sequence is a camera motion sequence that results in ambiguities in 3-D reconstruction when camera parameters are unknown. For example, only affine structures can be extracted from pure camera translational motion. Of particular relevance to the present disclosure is a determination that there is a two degree of freedom projective ambiguity for orbital motions, i.e., pure rotation about a fixed arbitrary axis. In other words, it has been demonstrated that there exists a 2 degree-of-freedom ambiguity in scaled Euclidean reconstruction.

[0012] There are three possible options in recovering scaled Euclidean structure from orbital motion:

- (1) fix two intrinsic camera parameters;
- (2) impose structural constraints, e.g., orthogonality, parallelism, known 3-D locations of fiducial points); or
- (3) get the "best" reconstruction without (1) or (2) as listed above.

SUMMARY OF THE INVENTION

[0013] The present invention provides a method for obtaining either a full or quasi-Euclidean reconstruction of an object image from a specific sequence of two dimensional images taken by an uncalibrated camera. In a closed-loop, the first image in the loop is positionally adjacent to the last image. The exact orientation of

the camera with respect to the object is not necessarily known. In one specific type of closed-loop sequence, the object (or camera) undergoes a complete orbital motion. Here, orbital motion means a pure rotation about a fixed arbitrary axis. Alternatively, the camera could move in a rectangular manner around the object while taking the closed-loop sequence of images.

[0014] The present method by-passes the usual intermediate stages of projective or affine reconstruction to recover a direct 3-D structure from feature point correspondences. The feature point correspondences are obtained from a two-stage bi-directional tracking process. 3-D reconstruction is done by applying the iterative Levenberg-Margardt algorithm to minimize error between actual point tracks and projected point tracks.

[0015] The invention in its broad form resides in a computerized method of reconstructing a three dimensional configuration from a sequence of two dimensional images, as recited in claim 1.

[0016] Although a preferred embodiment of the invention describes the specific case involving complete orbital motion, the invention envisages closed-loop motions of any type. During initialization of the method, the tilt of the axis of, for example, rotation is estimated by fitting ellipses on feature point tracks, and assuming equal angular rotation between frames. The feature tracks are determined by registering adjacent images, followed by feature point selection, and locally tracking the feature points. Since the images are arranged in a closed-loop, the tracking can be performed bi-directionally.

[0017] Once the point features have been tracked, an objective function is minimized to extract structure and motion parameters. The minimization is first performed by holding estimated skew and aspect ratio parameters constant, and second by letting the parameters vary until the results of the first and second minimization converge. Convergence is accelerated by adopting an object-centered structure-centered representation. Using a closed-loop sequence of images, the method yields superior results than those obtained for a partial sequence.

[0018] The 3-D structure can be recovered from a relatively sparse sequence of images where the camera intrinsic parameters, such as global camera tilt, local axis, amount of local movement between successive frames, and the 3-D positions associated with the tracked point features are all unknown. As an advantage of the method, the set-up is extremely simple using a camera, tripod stand, and a turntable. In addition, the global tilt angle, local rotation angles, and quasi-Euclidean structure can be extracted simultaneously. Furthermore, calibration of the camera is not required. The key idea is the initialization of the method based on the type of motion. In contrast with prior art methods described above, recovery of intermediate affine or projective structure is not necessary. The method exhibits fast convergence and is stable.

BRIEF DESCRIPTION OF THE DRAWINGS

[0019] A more detailed understanding of the invention may be had from the following description of a preferred embodiment, given by way of example and to be understood in conjunction with the accompanying drawing wherein:

Figure 1 is a block diagram of a system which uses the reconstruction method according to a preferred embodiment of the invention; and

Figures 2 is a flow diagram of global image registration and local feature tracking processes;

Figure 3 is a flow diagram of tilt and motion estimation processes; and

Figure 4 is a flow diagram of a process to minimize an objective function.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

System Overview

[0020] Figure 1 shows a computer system 100 connected to a camera 10, a display device 20, a keyboard 30, and a mouse 40. The system 100 includes one or more processors (P) 110, memories (M) 120, input/output interfaces (I/O) 130 connected to each other by bus lines 140.

[0021] During operation of the system 100, a three-dimensional (3-D) structure 12 makes a complete orbital motion, e.g., at least 360°. Herein, orbital motion means pure rotation about a fixed arbitrary axis 14. The angle of tilt 15 of the axis 14 with respect to the camera 10 is unknown. The structure 12 has sufficient surface texture to permit inter-frame feature tracking. Although the preferred embodiment is described using orbital motion, the invention can also be worked using other types of closed-loop motion.

[0022] While the structure 12 is rotating, the camera 10 measures light intensity and/or color values of the surface features as a closed-loop sequence of images. A closed-loop is defined herein as a sequence of images where the first and last image in the sequence are positionally adjacent. The camera 10 does not have significant radial distortion, or the radial distortion factor is known so that the images can first be corrected. Each image is an instantaneous planar (2-D) view of the structure 12. Typically, there are about 30 to 50 images in a sequence, that is, the sequence is relatively sparse. For some objects, the 3-D structure can be recovered from fewer frames. In general, better results are obtained with a larger number of frames.

[0023] The sequence of images is stored in the memories 120 of the system 100 as digitized frames 121. The first frame 1 is taken at a pose which is positionally adjacent to the last frame N. Typically, each frame 121 includes a plurality of pixels 122 arranged in a regular

pattern. The pixels 122 encode the measured light intensity/color values representative of the structure 14 at fixed angular poses. It should be noted that the same effect can be achieved by moving the camera 10 in a completely arbitrary orbital motion about a stationary structure. Other types of closed-loop motions are also applicable.

[0024] Computerized processes 123, described in greater detail below, are also stored in the memories 120 as instructions and data of software programs written, for example, in the C language. The programs operate under the control of an operating system, for example, UNIX®.

[0025] The instructions of the programs manipulate the pixel data of the frames 121 to perform a quasi-Euclidean (3-D) reconstruction of the object from the 2-D frames. A quasi-Euclidean reconstruction is close to, but not necessarily an exact scaled Euclidean construction. An exact scaled Euclidean reconstruction is not possible because of the 2 degree-of-freedom ambiguity inherent in orbital motion, as stated in the background section. Euclidean reconstruction would be possible in the case of general camera motion.

[0026] The 3-D shape of the structure 12 can be recovered from the relatively sparse sequence of images even though the camera intrinsic parameters, global camera tilt, local rotation axis, amount of local rotation between successive frames, and the 3-D positions associated with the tracked surface point features are all unknown.

[0027] A reconstructed 3-D object 21 can be viewed on the display device 20 by user commands initiated by the keyboard 30 and/or mouse 40. Because the 3-D shape is substantially recovered, the reconstructed object 21 can be viewed from any angle, even though the object was imaged with a fixed camera location, or in the case where the camera moves orbitally, the object is held in a fixed position.

[0028] Initialization of the preferred method is based on estimating the tilt 15 of the axis 14. In the case of orbital motion, the angle of tilt can be determined by fitting ellipses on feature point tracks. In addition, as an initial estimation, equal angular rotation between frames is first assumed. Then, bypassing the usual intermediate stages of projective or affine reconstruction as in the prior art, the method recovers 3-D structure directly from feature point correspondences obtained from a two-stage bi-directional tracking process. The 3-D reconstruction as disclosed herein is done by applying the iterative Levenberg-Marquardt algorithm to minimize error between actual point tracks and projected point tracks. Convergence is accelerated by adopting an object-centered representation.

Global Image Registration and Local Bi-Directional Feature Tracking

[0029] During initialization 201 as shown in Figure 2,

global image registration and local feature tracking are performed. Adjacent images are first registered in step 210. Image registration 210 determines the general image (optical) flow between successive frames in the sequence. Image registration compares at least two images to discern features (represented by pixel values) that are similar and different between the images.

[0030] Preferably, a pair-wise spline-based registration technique as described in U.S. Patent No. 5,611,000 "Spline-Based Image Registration" issued to Szeliski on 23 March 11, 1997 is used. In the spline-based technique, the image flow can generally be represented with a relatively small number of spline control nodes or vertices compared with the number of pixels. For each frame, feature points with high local texture, as indicated by their minimum eigenvalue of the local Hessian, are automatically selected for feature tracking.

[0031] Concurrently, a background mask is generated in step 212. The mask is used in step 214 to remove regions that do not change significantly in light intensity values. This is effective in removing a static background and low structure texture areas. After masking, the remaining features are locally tracked in step 216.

[0032] The tracking proceeds by way of bi-directional tracking. Bi-directional tracking involves comparing a frame with both adjacent frames. The selected feature points in each frame are individually and locally tracked using the flow estimates from the image registration step 210.

[0033] Here, the local tracking takes place independently in both (forward and backward) directions over multiple frames of the closed-loop. Because the frames form a closed-loop, every frame will have two adjacent frames. Multi-frame local tracking can be performed as described by Shi et al, in "Good Features to Track," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 593-600, June 1994.

[0034] These steps incorporate the advantages of both global spline-based registration and local tracking techniques. Although the global spline-based registration technique is able to track relatively significant motion, it is not able to deal with motion discontinuity very well due to its implicit smoothness constraint. On the other hand, the local tracking technique performs very poorly for significant motion but very well within the vicinity of the true track position.

[0035] In addition, because surface feature points of the structure 12 are tracked independently for local tracking, motion discontinuities resulting from the sparse sequence of images are handled correctly. Also, because the spline-based registration technique yields a reasonably good estimate of motion, the local tracking technique can then make use of this estimate to improve on the new track image position, especially when the feature position is within the vicinity of a motion discontinuity. By combining the methods as disclosed herein, better results can be obtained.

Tracking Termination

[0036] The tracking stops when any one of the following termination conditions occurs: each track continuously moves in one direction (220), the RMS pixel difference error (ε) is less than a predetermined threshold (222), for example, fifteen, or the minimum eigenvalue is above a predetermined threshold (224), for example five-hundred.

[0037] The problem with the first termination condition is that a complete track for some selected points may not be possible because that the points may not be observable in all the frames due to occlusions. However, the need to reject random noise is greater. Having a few more redundant points is a relatively small price to pay to reduce errors due to random noise. The tracks are post-processed to remove tracks that are deemed too short, for example, a minimum track length of 3 (image frames) can be imposed.

Tilt and Motion Determination

[0038] Next, as shown in Figure 3, tilt and motion are determined. In step 310, ellipses are fitted to the tracks determined above. From the orientation and shape of the ellipses, the tilt can be estimated. The estimated tilt, based on experimental results, is usually within 15° of the actual tilt.

[0039] The initialization of structure from motion recovery model for each frame 121 is made simpler by the fact that the motion of the structure 14 is, for example, a rotation about a unique, although unknown axis 14. If the camera 10 measures N images in the closed-loop sequence, then the local rotation angle about the z-axis 14 associated with image j is initialized to 360°/N in step 330. Errors in local motion estimates can be corrected as the images are further processed.

[0040] The camera factor and perspective distortion factor are initialized to values 1.0 and 0.001 respectively in step 340. Changes in these two values up to about an order of magnitude are well tolerated.

General 3-D Structure Recovery From Motion

[0041] According to the preferred embodiment of the invention, the three-dimensional shape of the object 12 involves the recovery of a set of 3-D structure parameters p_i , and time-varying motion parameters T_j from a set of measured image features u_{ij} as represented by the measured light intensity values, e.g., the pixels 122. A general formulation for linking 2-D image feature location u_{ij} of a frame j to its 3-D position p_i , where i is the track index, is:

$$u_{ij} = P(T_j^{(k)} \dots T_j^{(1)} p_i) \quad [Eq 1]$$

where the perspective projection transformation P() is applied to a cascaded series of rigid transformation $T_j^{(k)}$. Each transformation in turn can be defined by:

$$T_j^{(k)} x = R_j^{(k)} x + t_j^{(k)} \quad [Eq 2]$$

where R represents a rotation matrix and t is a translation applied after the rotation.

[0042] Within each of the cascaded transforms, the motion parameters may be time-varying (the j subscript is present) or fixed (the subscript is dropped). In the preferred embodiment of the invention, the transformation associated with the (horizontal) orbital motion can be expressed as:

$$u_{ij} = P(R_{x\tau} R_{z\theta} p_i + t) \quad [Eq 3]$$

where $R_{x\tau}$, $R_{z\theta}$ represent rotation about the x-axis by τ , and about the z-axis by θ respectively. It is assumed that there is negligible cyclotorsion, i.e., rotation of the camera 10 about the viewing axis.

[0043] The general camera-centered perspective projection equation can be formulated as:

$$\begin{pmatrix} u \\ v \end{pmatrix} = P \begin{pmatrix} x \\ y \\ z \end{pmatrix} \equiv \begin{pmatrix} \frac{sx + \eta\sigma y}{1 + \eta z} + u_0 \\ \frac{rsy}{1 + \eta z} + v_0 \end{pmatrix} \quad [Eq 4]$$

where s is the scale factor of the camera 10, r is the image aspect ratio, σ is the image skew, η is the perspective distortion factor, and (u_0, v_0) is the principal point in the image. The principal point is where the optical axis of the camera intersects the image plane.

Least-squares Minimization

[0044] The well known Levenberg-Marquardt algorithm can be used to solve for the structure and motion parameters. The merit or objective function to be minimized is:

$$C(a) = \sum_i \sum_j c_{ij} |u_{ij} - f(a_{ij})|^2 \quad [Eq 5]$$

where $f()$ is given in Eq 1, and a_{ij} consists of the unknown camera motion and the object structure parameters as given by:

$$a_{ij} = (p_i^T, m_j^T, m_g^T)^T \quad [Eq 6]$$

[0045] The vector a_{ij} contains all of the unknown structure and motion parameters, including the 3-D points p_i , the time-dependent motion parameters m_j , and the global motion/calibration parameters m_g . The weight c_{ij} in Eq 6 describes the confidence in the measured pixel values u_{ij} , and is normally set to the inverse variance σ_{ij}^{-2} .

Two-stage Approach

[0046] It is well known that there is a 2 degree-of-freedom ambiguity in scaled Euclidean reconstruction for orbital motion. A simplistic solution would set two of the intrinsic camera parameters to a constant, either as an assumption, or through camera calibration. A good choice, perhaps, would be to fix the image skew factor σ to 0, and the aspect ratio r to 1. However, in order to obtain better results, here the parameters may be kept unknown.

[0047] As shown in Figure 4, the preferred embodiment of the invention uses a two-stage approach to acquire the quasi-Euclidean structure.

First Stage

[0048] In step 410 of the first stage, fix the skew and the principle point to a constant value, e.g., $\sigma = u_0 = v_0 = 0$. The aspect ratio r is held constant at 1. Subsequently, apply the Levenberg-Marquardt algorithm in step 420 until terminating conditions are met. The termination conditions are either: the number of iterations exceeds a threshold, for example 150, or the improvement in the objective function is too small (10^{-6}), whichever comes first.

Second Stage

[0049] In step 430 of stage 2, let the parameters σ , u_0 , v_0 , and r vary, and continue with the Levenberg-Marquardt algorithm 340, with the iteration number reset to 0, until the same termination conditions are again met.

[0050] In general, using this two-stage approach

results in better convergence towards lower image projection errors and better structure and motion recovery.

[0051] As an advantage, a closed-loop sequence of images, where the first and last frames are positionally adjacent, yields a better reconstruction of the structure due to the extra constraints for feature points between the adjacent frames and their vicinities. If an incomplete rotation is used, as in most prior art methods, the errors in estimating the rotation angles are all biased in one direction due to the single "open" chain of constraints, causing a pinched appearance. In comparison, the reconstructed 3-D points of the same structure under complete rotation do not result in the same pinched appearance.

[0052] Note that in the case of general closed-loop camera motion, the first stage may be unnecessary. This is because there is no ambiguity in Euclidean reconstruction for arbitrary motion, in contrast with the two degrees of freedom ambiguity for orbital motion.

SUMMARY

[0053] Despite reconstruction ambiguities that exist for orbital motion, experimental results for synthetic and real 3-D structures show that reasonable quasi-Euclidean reconstruction can be done. This is because in practice, the camera intrinsic parameters of the principal point (u_0 , v_0), image skew σ , and aspect ratio r do not usually deviate significantly from normally assumed values of (0, 0), 0, and 1 respectively.

[0054] The preferred method operates according to a preference of maximizing knowledge of the camera to directly reconstructing 3-D shape, rather than converting to intermediate projective or affine representations of the structure. With a complete rotation of the structure, or an equivalent complete camera orbital motion which spans a full 360° , better results can be obtained. The better results are due to the additional constraints available in connection with the point tracks that bridges the first and last frames of the closed loop image sequence.

[0055] Described is a completely automatic method of recovering quasi-Euclidean structure from unknown by complete structure rotation, or equivalently, camera orbital motion. The preferred method starts with two-stage bi-directional tracking, followed by the application of iterative Levenberg-Marquardt minimization of feature point error to recover structure and rotational motion simultaneously. Apart from the knowledge that the images form a closed-loop, no other camera parameters are known.

[0056] The method minimizes the error between the projected image feature position and measured feature position directly and locally. In contrast with prior art techniques, no intermediate affine or projective reconstruction is done. The method usually converges toward the correct solution due to the assumption of complete structure rotation, which simplifies initialization. In addition,

tion, because the motion is rotational, a good initial estimate of the camera tilt can be made by fitting ellipses on the recovered tracks.

[0057] Having described a preferred embodiment of the invention, it will now become apparent to one skilled in the art that other embodiments incorporating its concepts may be used. It is felt therefore, that this invention should not be limited to the disclosed embodiment.

Claims

1. A computerized method for reconstructing a three-dimensional configuration of a structure from a sequence of two-dimensional images, comprising the steps of:

measuring the structure with a camera to generate a closed-loop sequence of images;
selecting feature tracking points in the closed-loop sequence of images;
estimating a relative orientation between the structure and the camera;
fixing a set of parameters while minimizing an objective function to extract first structure and motion parameters; and varying the set of parameters while minimizing the objective function to extract second structure and motion parameters until the second structure and motion parameters are substantially identical as the first structure and motion parameters.

2. The method of claim 1 wherein the closed-loop sequence of images is sparse, wherein the closed-loop sequence of images includes less than fifty images.
3. The method of claim 1 wherein intrinsic parameters of the camera are unknown.
4. The method of claim 1 wherein a relative displacement between successive images of the closed-loop sequence is unknown.
5. The method of claim 1 wherein the feature tracking points are unknown.
6. The method of claim 1 further including the step of estimating a local rotation angle from the number of images in the closed-loop sequence.
7. The method of claim 1 further including the step of masking static features in the closed-loop sequence of images.
8. The method of claim 1 further including the steps of:

registering adjacent pairs of images in the

closed-loop sequence of images; and
selecting the feature tracking points from the registered images.

9. The method of claim 8 wherein the image registration is spline-based.
10. The method of claim 1 wherein the feature tracking points are automatically selected.
11. The method of claim 1 wherein the selected feature tracking points are locally tracked until a termination condition is met.
12. The method of claim 1 wherein the objective function is minimized using a statistical optimal technique.
13. The method of claim 12 wherein the statistically optimal technique applies the iterative Levenberg-Marquardt algorithm.
14. The method of claim 1 wherein the closed-loop sequence of images is measured while the structure undergoes a complete orbital motion, wherein the axis of rotation with respect to the camera is unknown.
15. The method of claim 14 wherein ellipses are fitted to the selected feature tracking points, wherein the angle of tilt of the axis of rotation is estimated from the orientation and shape of the fitted ellipses.
16. The method of claim 1 wherein a first and last images measured in the sequence are positionally adjacent to form the closed-loop.

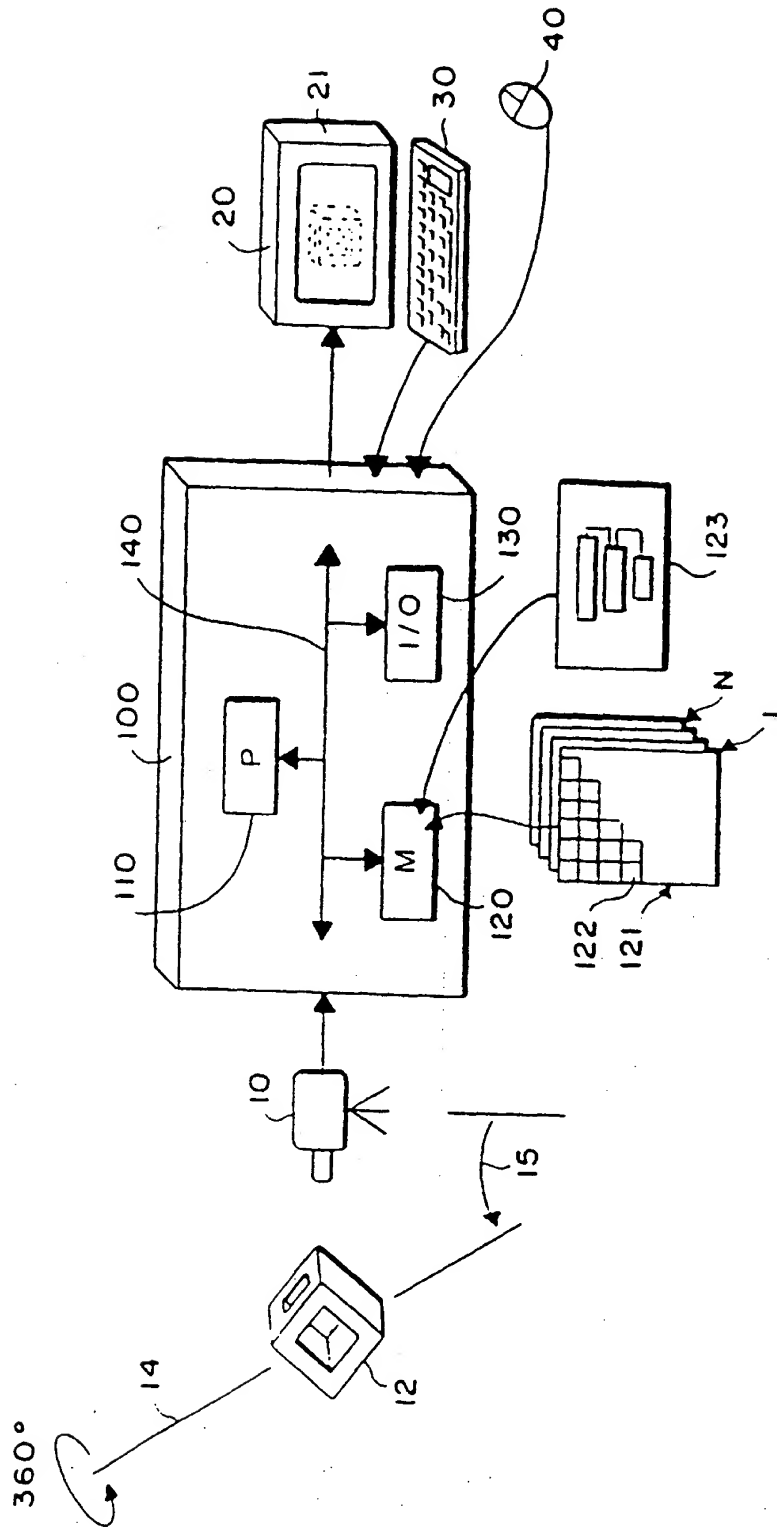


FIG. 1

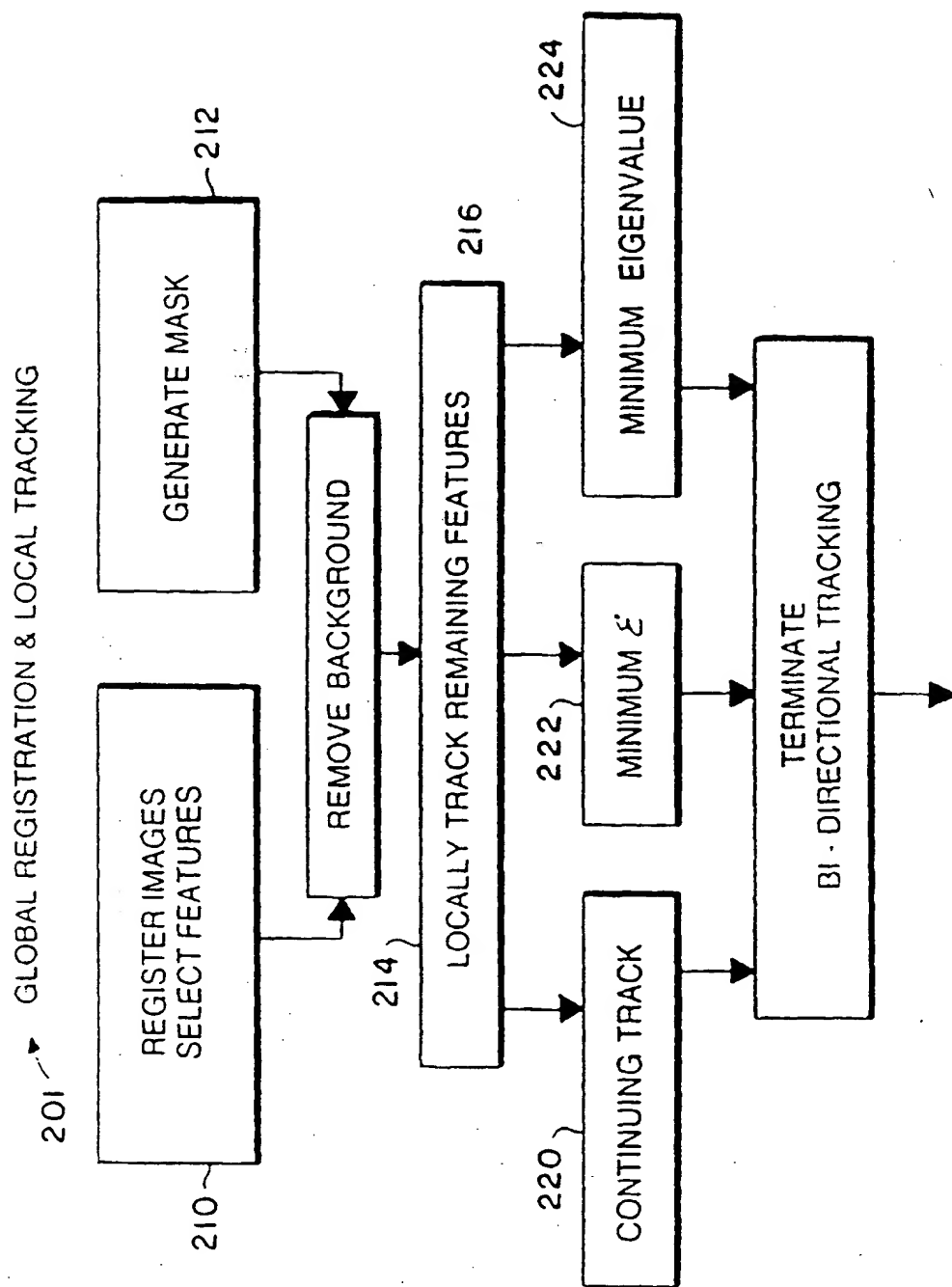


FIG. 2

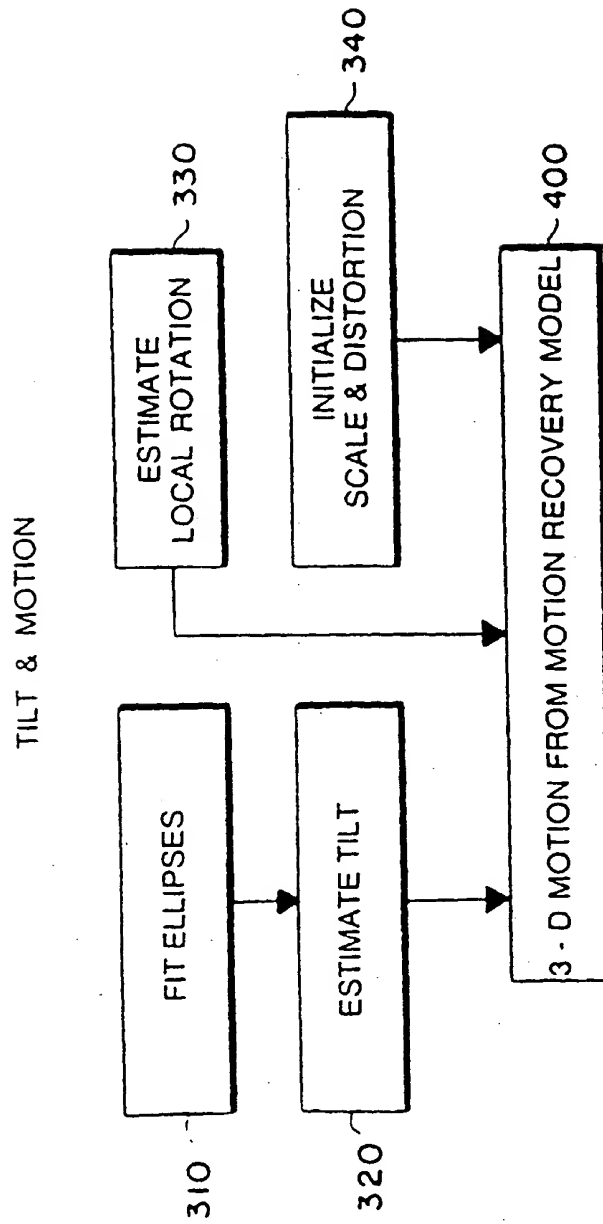


FIG. 3

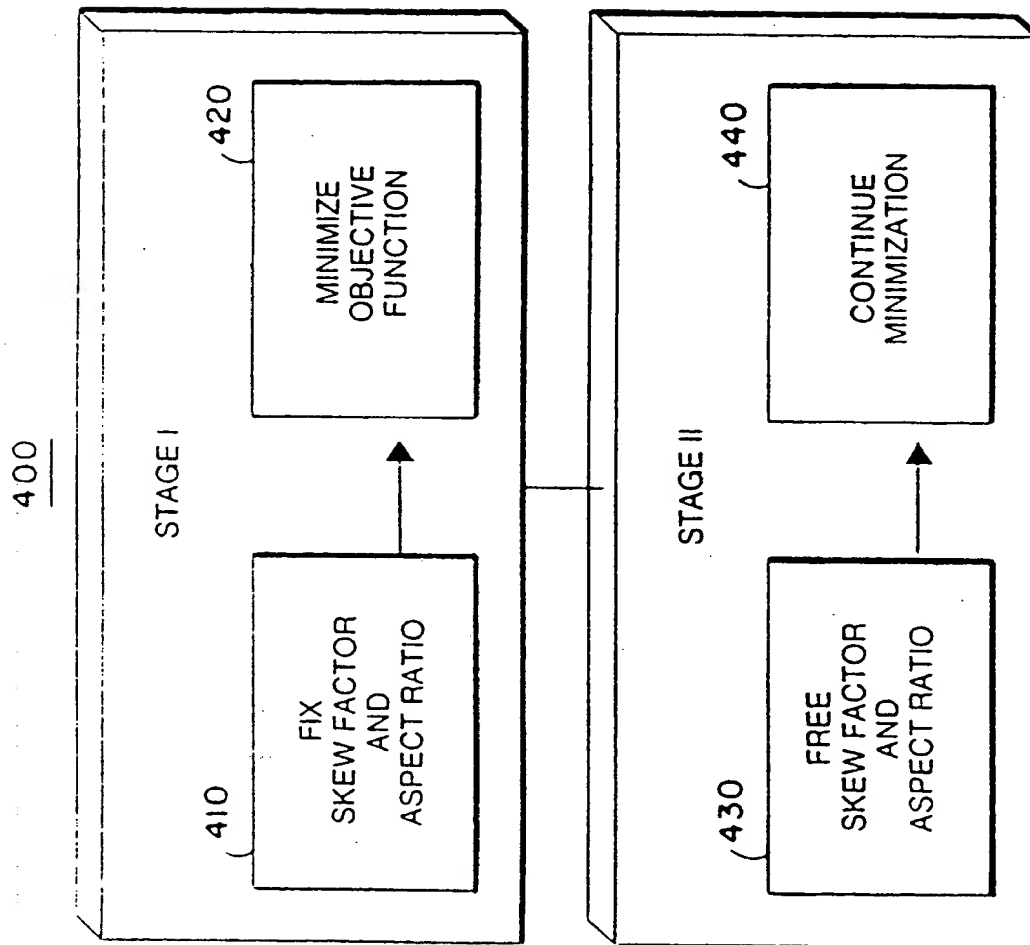


FIG. 4

(19)



Europäisches Patentamt

European Patent Office

Office européen des brevets



(11)

EP 0 895 192 A3

(12)

EUROPEAN PATENT APPLICATION

(88) Date of publication A3:
18.04.2001 Bulletin 2001/16

(51) Int. Cl.⁷: G06T 7/00, G06T 11/00

(43) Date of publication A2:
03.02.1999 Bulletin 1999/05

(21) Application number: 98114093.2

(22) Date of filing: 28.07.1998

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(72) Inventor: Kang, Sing Bing
Cambridge, MA 02138 (US)

(30) Priority: 28.07.1997 US 901391

(74) Representative:
Charig, Raymond Julian
Eric Potter Clarkson,
Park View House,
58 The Ropewalk
Nottingham NG1 5DD (GB)

(71) Applicant:
Compaq Computer Corporation
Houston Texas 77070 (US)

(54) **Method for reconstructing a three-dimensional object from a closed-loop sequence of images**

(57) In a computerized method, the three-dimensional structure of an object is recovered from a closed-loop sequence of two-dimensional images taken by a camera undergoing some arbitrary motion. In one type of motion, the camera is held fixed, while the object completes a full 360° rotation about an arbitrary axis. Alternatively, the camera can make a complete rotation about the object. In the sequence of images, feature tracking points are selected using pair-wise image registration. Ellipses are fitted to the feature tracking points

to estimate the tilt of the axis of rotation. A set of variables are set to fixed values while minimizing an image-based objective function to extract a set of first structure and motion parameters. Then the set of variables freed while minimizing of the objective function continues to extract a second set of structure and motion parameters that are substantially the same as the first set of structure and motion parameters.

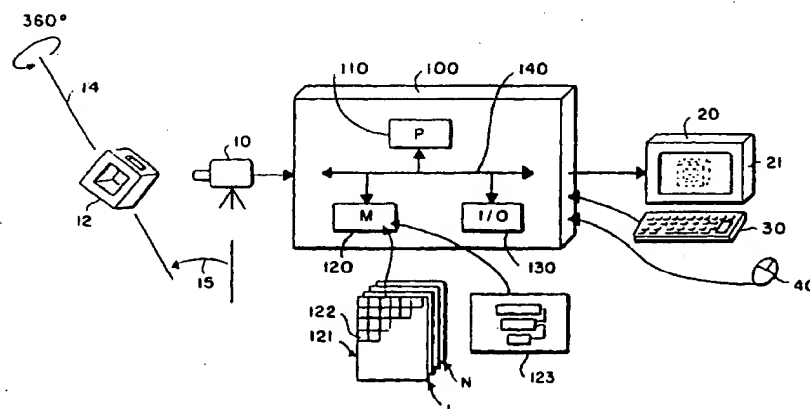


FIG. 1

EP 0 895 192 A3



European Patent
Office

EUROPEAN SEARCH REPORT

Application Number
EP 98 11 4093

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.8)
A	SZELISKI R ET AL: "A parallel feature tracker for extended image sequences" PROCEEDINGS INTERNATIONAL SYMPOSIUM ON COMPUTER VISION (CAT. NO.95TB100006), ISCV, CORAL GABLES, FL, USA, 21-23 NOV. 1995, pages 241-246, XP002160564 1995, Los Alamitos, CA, USA, IEEE Comput. Soc. Press, USA ISBN: 0-8186-7190-4 * sections 3 to 5 *	1-16	G06T7/00 G06T11/00
A	SZELISKI R ET AL: "Recovering 3D shape and motion from image streams using nonlinear least squares" JOURNAL OF VISUAL COMMUNICATION AND IMAGE REPRESENTATION, MARCH 1994, USA, vol. 5, no. 1, pages 10-28, XP000983099 ISSN: 1047-3203 * sections 1 and 7.1 *	1-16	
A	HEYDEN A ET AL: "Euclidean reconstruction from image sequences with varying and unknown focal length and principal point" PROCEEDINGS. 1997 IEEE COMPUTER SOCIETY CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CAT. NO.97CB36082), SAN JUAN, PUERTO RICO, 17-19 JUNE 1997, pages 438-443, XP002160565 1997, Los Alamitos, CA, USA, IEEE Comput. Soc, USA ISBN: 0-8186-7822-4 * section 1 *	1-16	TECHNICAL FIELDS SEARCHED (Int.Cl.8) G06T
The present search report has been drawn up for all claims			
Place of search BERLIN		Date of completion of the search 19 February 2001	Examiner Bouchaala, N
<p>CATEGORY OF CITED DOCUMENTS</p> <p>X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document</p> <p>T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document</p>			

EPO FORM 1503 03 82 (P04C01)



European Patent
Office

EUROPEAN SEARCH REPORT

Application Number
EP 98 11 4093

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.CI.8)
A	<p>VERRI A ET AL: "Aspect-based object recognition with size functions" PROCEEDINGS OF 13TH INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION, VIENNA, AUSTRIA, 25-29 AUG. 1996, pages 682-686 vol.1, XP002160566 1996, Los Alamitos, CA, USA, IEEE Comput. Soc. Press, USA ISBN: 0-8186-7282-X * section 2, paragraphs 3 and 4 *</p>	2-6	
A	<p>GHOSH B K ET AL: "Calibration free visually controlled manipulation of parts in a robotic manufacturing workcell" PROCEEDINGS. 1996 IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION (CAT. NO.96CH35857), MINNEAPOLIS, MN, USA, 22-28 APRIL 1996, pages 3197-3202 vol.4, XP002160567 1996, New York, NY, USA, IEEE, USA ISBN: 0-7803-2988-0 * abstract * * section 2, paragraphs 4 and 5 *</p>	14-16	
The present search report has been drawn up for all claims			<p>TECHNICAL FIELDS SEARCHED (Int.CI.8)</p>
Place of search		Date of completion of the search	Examiner
BERLIN		19 February 2001	Bouchaâla, N
<p>CATEGORY OF CITED DOCUMENTS</p> <p>X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document</p> <p>T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document</p>			

EPO FORM 1503 03/82 (P04C01)

This Page Blank (uspto)

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

This Page Blank (uspto)